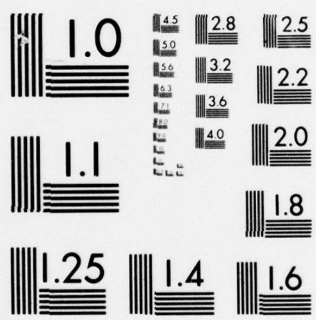
AD-A054 553 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
THE USE OF STOCHASTIC PROGRAMMING FOR THE SOLUTION OF SOME PROB--ETC(U)
MAR 78 A PREKOPA DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1834 NL

| OF |

AD
A054553

END
DATE
FILMED
6-78
DDC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

MRC Technical Summary Report #1834

THE USE OF STOCHASTIC PROGRAMMING FOR THE SOLUTION OF SOME PROBLEMS IN STATISTICS AND PROBABILITY

András Prékopa

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

March 1978

(Received January 31, 1978)

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

THE USE OF STOCHASTIC PROGRAMMING FOR THE SOLUTION OF SOME PROBLEMS IN STATISTICS AND PROBABILITY

András Prékopa

Technical Summary Report #1834
March 1978

ABSTRACT

The applicability of known stochastic programming models and methods for the solution of problems in classical statistics and probability is shown by a number of examples. These concern testing of hypotheses, constructing of tolerance regions, planning of optimal sampling and the Moran model for the dam.

AMS (MOS) Subject Classifications: 90C15, 90C25, 62D05, 62H15, 60A99

Key Words: Chance constrained programming, Statistical decision, Testing hypotheses, Dam problem

Work Unit Number 5 (Mathematical Programming and Operations Research)

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

1473

SIGNIFICANCE AND EXPLANATION

Mathematical programming deals with optimization under constraints, e.g. minimization of costs in manufacturing operations when sources of supply, machines and manpower are limited. Stochastic programming deals with mathematical programming problems in which there is uncertainty associated with the variables and with the constraints in the optimization problem. Thus stochastic programming lies on the borderline between mathematical programming and statistics. However, the connection between the development of these two sciences is not strong enough.

The purpose of this paper is to show how some stochastic programming methods (developed by the author) can be applied in classical problems of statistics. Models are formulated for a) construction of statistical tests, b) construction of tolerance regions, c) optimum allocation in surveys, d) the dam problem of MORAN. The solution of the above problems uses nonlinear programming combined with simulation. Evidence concerning effective solvability of such problems is given in other referenced papers.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

THE USE OF STOCHASTIC PROGRAMMING FOR THE SOLUTION OF SOME PROBLEMS *IN STATISTICS AND PROBABILITY*

András Prékopa

1. *Introduction*

Stochastic programming is a branch of mathematical programming because it deals with handling of mathematical programming problems where some of the parameters are random variables. At the same time we may consider it to be part or extension of statistics [4] because problems concerning optimization of stochastic systems in general can be considered to belong to statistics in the wide sense. We say in the "wide sense" because in many textbooks on statistics the definition of this science includes the possibility of gaining information by experimentation to an extent depending on the statistician (see e.g. the Introduction of [14]). It is, however, not always the case concerning the stochastic systems stochastic programming deals with. If e.g. we formulate reservoir system design models using past hydrological data then it is not possible to perform further experimentation unless we postpone the date of the building of the reservoirs. Some statisticians do not require that the possibility of getting information by further experimentation should be characteristic for the statistical methods. Accepting this, it is in fact appropriate to say that stochastic programming is an "extension of statistics".

For about twenty years, however, stochastic programming has had a development almost independent of the parallel development of the statistical methods. Hence it is reasonable to show that some of the stochastic programming models and methods provide useful tools for the solution of known problems in classical statistics and probability. The purpose of this paper is to show this possibility by a number of examples.

The possibility of applying stochastic programming in statistics and probability is easy to understand. In fact the decision principles of

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

stochastic programming contain ideas taken from classical statistical theories. Prescribing lower or upper bound for probabilities and find optimal strategy is a general way of thinking in testing hypothesis. To decide, then observe and to decide again, appears to be the general decision scheme in the decision theory of Wald and others. To find optimal strategy by minimizing an expectation depending on some decision variable(s) already appears in the work of D. Bernoulli concerning the solution of the Petersburg problem [16]. However, the fact that similar principles can be applied and how they can be applied for large scale problems described by mathematical programming problems where some of the parameters are random, was by no means simple to discover. Because not only the formal application of the statistical principles was necessary to carry out but the new problems turning out needed to be numerically solvable mathematical programming problems and these must have resulted in good applications.

The problems we formulate in this paper are of "chance constrained" or in other words "probabilistic constrained" type. This type of stochastic programming model has been introduced by Charnes, Cooper and Symonds [3] (the paper was presented at the Econometric Society meeting in Washington, D.C., 1953). In our models, however, we frequently impose constraint on the joint probability for the occurrence of random events which is important in particular in applications in statistics and probability from the point of view of model construction. At the same time this makes the numerical solution of the problem more complicated.

Since extensive application has been done by the author and his collaborators concerning such stochastic programming problems, evidence exists that problems of the type we are going to formulate are numerically solvable at least in moderate sizes.

The problems we are going to formulate are of the following type

(1.1) minimize $f(x)$ subject to

$$h_0(x) = P(g_1(x,\xi) \geqq 0,\dots,g_r(x,\xi) \geqq 0) \geqq p ,$$

$$h_i(x) \geqq 0, \quad i = 1,\dots,m ,$$

where $x \in R^n$, $g_1,\dots,g_r$ are functions of the deterministic variable x and the random variable ξ further $h_1,\dots,h_m$ are functions of the deterministic variable x. Problems of the type (1.1) are solved by the combined use of nonlinear programming and simulation where the latter is used to determine function values which are probabilities and gradient values which can be expressed in terms of probabilities concerning h_0 and these probabilities belong to sets in higher dimensional spaces.

If Problem 1 is a convex programming problem i.e. the set of feasible solutions is convex and $f(x)$ is a convex function, then when solving the problem we can reach global optimum, otherwise we can only expect local optimum. To have a convex programming problem it is crucial to know that the constraint $h_0(x) \geqq p$ determines a convex set. General theorems ensure this property. For the reader's convenience we formulate here two theorems proved in [9], [10] and [1], [2], respectively. For further references see [13].

Theorem 1. If $g_1,\dots,g_r$ are concave functions in all variables in the entire space and ξ has a logarithmically concave[1] probability density function, then h_0 is a logarithmically concave function in the entire space of the variable x.

---

[1] A nonnegative point function h defined on a convex set A is said to be logarithmically concave (logconcave) if for every $x, y \in A$ and $0 < \lambda < 1$, we have $h(\lambda x + (1-\lambda)y) \geqq [h(x)]^{\lambda}[h(y)]^{1-\lambda}$.

Theorem 2. If $x \in R^n$, $\xi \in R^q$ and $g_1, \ldots, g_r$ are concave functions in R^{n+q}, further the probability density function f of the random vector ξ has the property that $f^{-1/n}$ is convex in R^q, then h_0 is a quasi-concave[2] function in R^n.

In some of the problems formulated in this paper we shall allow more than one probabilistic constraints.

[2]A function h defined on a convex set A is said to be quasi-concave if for every $x, y \in A$ and $0 < \lambda < 1$, we have $f(\lambda x + (1-\lambda)y) \geqq \min[f(x), f(y)]$.

2. Testing Statistical Hypothesis and Construction of Confidence Region

The simplest example that can be formulated to illustrate testing hypothesis is to test the probability distribution P_0 against the alternative P_1. For the sake of simplicity we assume that these are discrete distributions assigning positive probabilities only to positive integers and $P_0(k)$, $P_1(k)$ respectively, designate the probabilities belonging to the integer k. To construct a test of size less than or equal to the prescribed probability α and having maximum power is equivalent to finding a set of integers S such that

$$\begin{aligned} &\text{first kind error} = \sum_{k\in S} P_0(k) \leqq \alpha , \\ (2.1) \qquad &\\ &\text{power of the test} = \sum_{k\in S} P_1(k) \to \max . \end{aligned}$$

This problem can be formulated as a mathematical programming problem

$$\begin{aligned} &\text{maximize} \quad \sum_{k\in S} P_1(k)x_k \\ (2.2) \qquad &\text{subject to} \quad \sum_{k\in S} P_0(k)x_k \leqq \alpha , \\ &x_k = 0 \text{ or } 1 \text{ all } k . \end{aligned}$$

The optimal set S is then given by

$$S = \{k \mid x_k = 1\} .$$

Problem (2.2) is known in the operations research as the knapsack problem and has a simple solution mentioned already in [7, p. 64]. We have mentioned the above problem only in order to present a starting point for the formulation of more sophisticated tests. We introduce the following notations

H: Set of probability distributions representing a hypothesis.

K: Set of probability distributions representing the alternative.

X: Vector valued random variable. On the basis of the observed value of X we reject or accept the hypothesis.

α: Prescribed upper bound for the first kind errors.

β: Variable to be maximized, its optimum value is the maximum value of the minimum power of the test.

S: Critical region.

The problem is to find S by solving the following problem

$$\text{maximize } \beta \text{ subject to}$$

$$\begin{aligned} &P(X \in S|F) \leqq \alpha \quad \text{for every} \quad F \in H\,, \\ &P(X \in S|F) \geqq \beta \quad \text{for every} \quad F \in K\,. \end{aligned} \tag{2.3}$$

In (2.3) on the left hand sides there stand probabilities of the event $X \in S$ using the probability distribution F for the random vector x.

Problem (2.3) is too general. We can hardly find a set S this way. We can, however, restrict the type of the set S and find one in the restricted category. To this end let us introduce the sets

$$S_k = \{x|L_k(x) \geqq b_k\}, \qquad k = 1,\dots,m$$

$$S = \bigcup_{k=1}^{m} S_k\,,$$

where $L_1,\dots,L_k$ are given linear forms of the variable $x \in R^n$ and $b_1,\dots,b_m$ are real parameters the values of which we want to determine by an optimization problem. The closure of the complementary of S is a convex polyhedron for every fixed values of the parameters $b_1,\dots,b_m$ and through the variation of the parameters only (some of) the faces are shifted (see Fig. 1). Let us assume that all probability distributions in H and K are continuous. This implies that each of them assigns 0 probability to any hyperplane. Knowing this, we can write

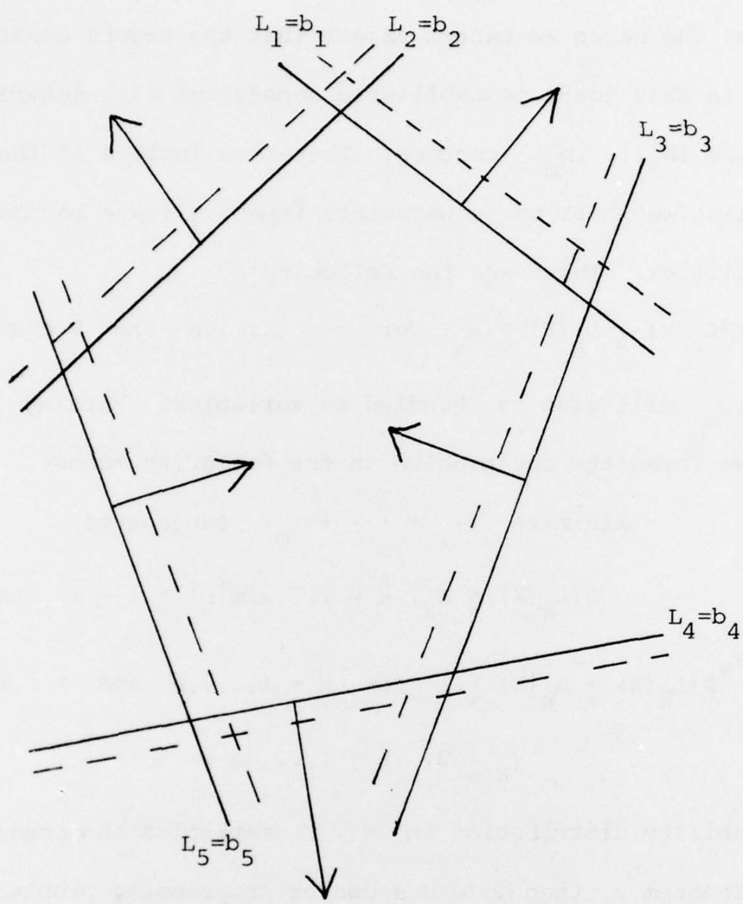


Figure 1

$$P(X \in S|F) = 1 - P(L_k(X) \leqq b_k,\ k = 1,\ldots,m|F)$$

for every $F \in H$.

This is the probability in the first constraint of Problem (2.3). The constraint itself can be written in the following manner:

$$(2.4) \qquad P(L_k(X) \leqq b_k,\ k = 1,\ldots,m|F) \geqq 1 - \alpha \quad \text{for} \quad F \in H .$$

If every $F \in H$ has a logconcave density or a density satisfying the condition of Theorem 2, then the set of vectors $(b_1,\dots,b_m)$ satisfying (4) is a convex set. This is ensured by Theorem 1 and 2, whichever applies.

In most of the cases we cannot expect that the second constraint in Problem (2.3) in this joint probabilistic constraint will determine a convex set of the $(b_1,\dots,b_m)$ vectors. Therefore instead of the single second constraint we shall use constraints imposing lower bounds for the single probabilities. These are the following

$$P(L_k(X) \geqq b_k | F) \geqq \beta_k \quad \text{for} \quad k = 1,\dots,m \quad \text{and} \quad F \in K, \tag{2.5}$$

where $\beta_1,\dots,\beta_m$ will also be handled as variables. Putting $\beta_k = e^{-\gamma_k}$, $k = 1,\dots,m$ we formulate our problem in the following manner

$$\begin{aligned} &\text{minimize} \quad (\gamma_1 + \cdots + \gamma_m) \quad \text{subject to} \\ &P(L_k(X) \leqq b_k,\ k = 1,\dots,m | F) \geqq 1 - \alpha \quad \text{for} \quad F \in H, \\ &e^{\gamma_k} P(L_k(X) \geqq b_k | F) \geqq 1 \quad \text{for} \quad k = 1,\dots,m \quad \text{and} \quad F \in K, \\ &\gamma_k \geqq 0, \quad k = 1,\dots,m. \end{aligned} \tag{2.6}$$

If every probability distribution in $H \cup K$ satisfies the condition of Theorem 1 or Theorem 2, then (2.6) is a convex programming problem. Minimizing

$$\sum_{k=1}^{m} e^{-\gamma_k}$$

instead of the sum of $\gamma_1,\dots,\gamma_m$, the convex programming character of the problem will not be disturbed.

Similar problem can be formulated for the construction of confidence region. In this case we use the first constraint of Problem (2.6), impose some other constraints on $b_1,\dots,b_m$ (these can be e.g. lower and upper bounds) and minimize a certain function of the variables $b_1,\dots,b_m$. Thus we are lead to the following problem:

$$\text{minimize } f(b) \text{ subject to}$$

$$(2.7) \qquad P(L_k(X) \leqq b_k,\ k = 1,\ldots,m \mid F) \geqq 1 - \alpha \quad \text{for} \quad F \in H ,$$

$$h_i(b) \geqq 0, \quad i = 1,\ldots,M .$$

If every F in H satisfies the condition of Theorem 1 or Theorem 2, then the first constraint determines a convex set of the b vectors. Then, if f is convex and $h_1,\ldots,h_M$ are concave or quasi-concave, Problem (2.7) will be a convex programming problem. The confidence region will be the intersection of the sets $S_1,\ldots,S_m$ using the optimal $b_1,\ldots,b_m$ produced by Problem (2.7). The above notion of a confidence region is slightly different from the conventional one. The reason is that we work with general sets of probability distributions and disregard the parameter. It is not difficult, however to specialize our problem formulation for the case when we have parameter and to derive this way confidence region in the conventional sense for the unknown parameter.

3. Construction of Tolerance Regions

Let F be the probability distribution of the random vector $X = (X_1,\dots,X_N)$. The random set S is a β-content tolerance region of confidence level γ if the following inequality holds

$$P(\int_S dF \geqq \beta) \geqq \gamma \ . \tag{3.1}$$

The random set S is constructed on the basis of a sample taken from the population F. The probability standing on the left hand side in (3.1) is called the coverage of S.

Instead of a general formulation of the construction of a tolerance region based on stochastic programming, we show the principal idea on a simple problem.

Suppose we want to construct tolerance region of the form $S = (0, K\bar{x})$ for an exponential distribution with unknown parameter λ where $\bar{x}$ is the empirical mean of a sample of size n taken from this population and K is a number to be determined so that it should be the smallest number satisfying (3.1) with given β and γ. In our case (3.1) can be written in the following manner

$$P(1 - e^{-\lambda K\bar{x}} \geqq \beta) \geqq \gamma \ . \tag{3.2}$$

Let us introduce the notation $\bar{y} = \lambda\bar{x}$. The probability distribution of $\bar{y}$ is independent of λ and $n\bar{y}$ has a standard gamma distribution with parameter n i.e. $n\bar{y}$ has the following probability density function:

$$f(z) = \frac{z^{n-1}e^{-z}}{(n-1)!} \quad \text{if} \quad z > 0$$

and $f(z) = 0$ otherwise. The inequality (3.2) can be rewritten as follows:

$$P(n\bar{y} \geqq \frac{n}{K} \log \frac{1}{1-\beta}) \geqq \gamma \ . \tag{3.3}$$

It will be more convenient to use the new variable $L = \frac{1}{K}$. Putting $K = \frac{1}{L}$ in (3.3), our problem will be to find the largest L satisfying (3.3).

Assume now that we have m independent exponentially distributed random variables with parameters $\lambda_1,\ldots,\lambda_m$, respectively. Let $\bar{x}_1,\ldots,\bar{x}_m$ denote the sampling expectations corresponding to independent samples of sizes $n_1,\ldots,n_m$ taken from these populations and $\bar{y}_i = \lambda \bar{x}_i$, $i = 1,\ldots,m$. For the construction of a tolerance region of the form ($\times$ means Cartasian product):

$$S = (0, \frac{1}{L_1}\bar{x}_1) \times (0, \frac{1}{L_2}\bar{x}_2) \times \cdots \times (0, \frac{1}{L_m}\bar{x}_m) \tag{3.4}$$

it is reasonable to choose the following decision principle

$$\text{maximize} \quad (L_1 + \cdots + L_m) \quad \text{subject to}$$

$$\prod_{i=1}^{m} P(n_i\bar{y}_i \geqq L_i n_i \log \frac{1}{1-\beta}) \geqq \gamma , \tag{3.5}$$

$$a_i \leqq L_i \leqq b_i, \quad i = 1,\ldots,m ,$$

where $a_1,\ldots,a_m$; $b_1,\ldots,b_m$ are prescribed bounds. Since a standard gamma density function with parameter greater than or equal to 1, is a logconcave point function, it follows from Theorem 1 that the constraining function i.e. the function standing on the left hand side of the first constraint of Problem (3.5), is a logconcave function of the variables $L_1,\ldots,L_m$. Thus (3.5) is a convex programming problem.

It is not necessary to restrict ourselves to the case of independent random variables. Consider e.g. the random variables

$$x_1 = \frac{1}{\lambda_1}(z_1 + z_2) ,$$

$$x_2 = \frac{1}{\lambda_2}(z_1 + z_3) ,$$

where λ_1, λ_2 are positive (unknown) constants and z_1, z_2, z_3 are standard gamma distributed random variables with known parameters $\vartheta_1, \vartheta_2, \vartheta_3$. If we take three independent samples concerning z_1, z_2, z_3, all of them of size n and use the notations $\bar{z}_1, \bar{z}_2, \bar{z}_3$ for the sample means, further $\bar{y}_1 = \lambda\bar{z}_1$, $\bar{y}_2 = \lambda\bar{z}_2$, $\bar{y}_3 = \lambda\bar{z}_3$, then we can formulate the following stochastic programming problem

$$\text{maximize} \quad (L_1 + L_2) \quad \text{subject to}$$

$$P\begin{pmatrix} n(\bar{y}_1 + \bar{y}_2) \geqq nL_1 \log \dfrac{1}{1-\beta} \\ \\ n(\bar{y}_1 + \bar{y}_3) \geqq nL_2 \log \dfrac{1}{1-\beta} \end{pmatrix} \geqq \gamma , \tag{3.6}$$

$$a_i \leqq L_i \leqq b_i, \quad i = 1,2 .$$

The random variables $n\bar{y}_1, n\bar{y}_2, n\bar{y}_3$ are independent and they have standard gamma distributions with parameters $n\vartheta_1, n\vartheta_2, n\vartheta_3$. If $n\vartheta_1 \geqq 1$, $n\vartheta_2 \geqq 1$, $n\vartheta_3 \geqq 1$, then $n\bar{y}_1, n\bar{y}_2, n\bar{y}_3$ have a logconcave joint density. Hence by Theorem 1, the constraining function in the first constraint of Problem (3.6) is a logconcave function of the variables L_1, L_2. Thus (3.6) is a convex programming problem. The tolerance region will be that special case of (3.4) where $m = 2$ and L_1, L_2 are the components of an optimal solution of Problem (3.6).

4. Optimum Allocation in Surveys

A well-known application of nonlinear programming is to find the number of elements in the different strata concerning stratified sampling from a finite population. First we formulate this problem that is a deterministic one, then using it as an underlying deterministic problem, we formulate a stochastic programming problem. Let us introduce the following notations:

L	Number of strata
N_h	Elements in stratum h
$N = \sum_{h=1}^{L} N_h$	Total number of elements in the population
n_h	Unknown number of elements to be chosen from stratum h
$X_h = \frac{1}{n_h} - \frac{1}{N_h}$,	
$W_h = Nh/N$	
$\bar{y}_j$	Estimate of the jth variable
r	Number of variables to be estimated
$S^2_{h_j}$	Variance of the jth variable in stratum h
V^2_j	Variance of the estimate $\bar{y}_j$
$a_{hj} = W^2_h S^2_{hj}$	
d_j	Prescribed numerical upper bound for V^2_j
C_h	Unit price of sampling from stratum h .

It is well-known that the variance V^2_j can be expressed in the following manner

$$V^2_j = \sum_{h=1}^{L} a_{hj} X_h, \quad j = 1,\ldots,r .$$

To find $n_1,\ldots,n_L$, we formulate a nonlinear programming problem. In order to have linear constraints, we prefer to use the variables $X_1,\ldots,X_L$. Since

$$n_h = 1/(X_h + \frac{1}{N_h}) ,$$

our problem reads as follows

$$\text{minimize} \quad \sum_{h=1}^{L} C_h/(X_h + \frac{1}{N_h}) \quad \text{subject to}$$

$$\sum_{h=1}^{L} a_{hj}X_h \leqq d_j, \qquad j = 1,\ldots,r ,$$

$$0 \leqq X_h \leqq 1 - \frac{1}{N_h}, \quad h = 1,\ldots,L .$$

In this problem the constraints are linear and the objective function to be minimized is convex. Hence (4.1) is a convex programming problem.

Assume now that within the strata we have such populations the variances of which are random variables. Then we can impose a probabilistic constraint on the first p constraints of Problem (4.1) and formulate the following new problem

$$\text{minimize} \quad \sum_{h=1}^{L} C_h(X_h + \frac{1}{N_h}) \quad \text{subject to}$$

$$P(\sum_{h=1}^{L} a_{hj}X_h \leqq d_j, \quad j = 1,\ldots,r) \geqq p$$

(4.2)

$$0 \leqq X_h \leqq 1 - \frac{1}{N_h}, \quad h = 1,\ldots,L .$$

We may take a small sample before and use the aposteriori distribution of the coefficients a_{ij} given the result of the small sample. In this case the structure of the problem (4.2) remains but we have new probability distribution for the random variables in the first constraint.

Problems of the type (4.2) are frequently nonconvex. Here the coefficients of the unknowns are random in the probabilistic constraint.

Some result concerning programming under probabilistic constraint with random technology matrix are presented in [5], [11]. According to these results still in many cases (4.2) will be a convex programming problem.

5. An Example Concerning Stochastic Processes

We consider the Moran model for the dam [8] and see how stochastic programming can improve this model.

Time will be subdivided into discrete periods and we number them by 1,2,... . Let K be the capacity of the dam. Assume that in the beginning of Period i an input occurs: out of a total input x_i that amount for which we have freeboard, fills up the reservoir to that extent and the remaining water overflows. After this, an output (release) occurs. We release an amount equal to M if at least that amount is available and we release the total amount from the dam if the available amount is smaller than M. Let z_i denote the water content of the reservoir at the end of Period i. The following recursive relation holds true

$$z_i = \max[\min(z_{i-1} + x_i, K) - M, 0], \quad i = 1,2,... \tag{5.1}$$

where z_0 is the initial water content of the reservoir. All demands will be met in the course of the first n periods if and only if the following relations hold:

$$z_i = \min(z_{i-1} + x_i, K) - M \geqq 0, \quad i = 1,...,n . \tag{5.2}$$

If $x_1, x_2, ...$ are independent, identically distributed random variables, then $z_1, z_2, ...$ form a Markov chain. Under mild conditions we have ergodicity [8] and using the stationary limit distribution, a reservoir capacity design principle can be formulated so that we put

$$P(\min(z_{i-1} + x_i, K) \geqq M) = p , \tag{5.3}$$

where p is a prescribed high probability.

We now drop the condition imposed on the random variables $x_1, x_2, ...$ and assume only that concerning a finite subsequence $x_1, ..., x_n$, the condition of Theorem 1 or Theorem 2 is fulfilled. We also drop the condition that the amount of water to be released is constant and does not depend on the period. Moreover, we shall introduce the water quantities $M_1, \ldots, M_n$ to be released in the subsequent periods as unknowns. In order to obtain the recursive relations for this case we only have to write M_i instead of M in (5.1). Writing M_i instead of M in (5.2) we obtain a necessary and sufficient condition that all demands can be met in the course of the n periods.

For the determination of the capacity K and $M_1, \ldots, M_n$ we formulate the following stochastic programming problem:

$$\text{minimize}[c(K) - c_1 M_1 - \cdots - c_n M_n] \text{ subject to}$$

$$P(\min(z_{i-1} + x_i, K) - M_i \geqq 0, \quad i = 1, \ldots, n) \geqq p, \tag{5.4}$$

$$0 \leqq K \leqq K_0$$

$$0 \leqq M_i \leqq M_{i0}, \quad i = 1, \ldots, n,$$

where p is a prescribed high probability, K_0, M_{i0}, $i = 1, \ldots, n$ are given constants, $c(K)$ is the building cost of the reservoir and $c_1, \ldots, c_n$ are the benefits of water units in the subsequent periods. We can e.g. assume that winter moisture fills up completely the reservoir thus Periods $1, \ldots, n$ are some few months in the spring, summer and fall. If the reservoir serves for 50 years, say, then the building cost should be subdivided into 50 equal parts in a discounted form and use only the first year part as $c(K)$ to have a right economic formulation of the problem. This is, however, not a central point of our present discussion.

Problem (5.4) is a convex programming problem if $c(K)$ is a convex function and the joint distribution of $x_1, \ldots, x_n$ satisfies the condition of Theorem 1 or Theorem 2.

For further reservoir system design models based on stochastic programming the reader is referred to [10], [12].

References

[1] C. Borell, "Convex set functions in d-space", Periodica Mathematica Hungarica 6 (1975) 111-136.

[2] H. J. Brascamp and E. H. Lieb, "On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for logconcave functions, and with an application to diffusion equation", Journal of Functional Analysis 22 (1976) 366-389.

[3] A. Charnes, W. W. Cooper and G. H. Symonds, "Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil production", Management Science 4 (1958) 236-263.

[4] A. Charnes, W. W. Cooper and M. J. L. Kirby, "Chance-constrained programming: an extension of statistical method", in [15], 391-402.

[5] R. Jagannathan and M. R. Rao, "A class of nonlinear chance-constrained programming models with joint constraints", Operations Research 21 (1973) 360-364.

[6] A. R. Kokan and S. Khan, "Optimum allocation in multivariate surveys: an analytical solution", Journal of the Royal Statistical Society B29 (1967) 115-125.

[7] E. L. Lehmann, Testing statistical hypotheses (Wiley, New York, 1959).

[8] P. A. P. Moran, The theory of storage (Methuen, London, 1959).

[9] A. Prékopa, "Logarithmic concave measures with application to stochastic programming", Acta Scientiarum Mathematicarum (Szeged) 32 (1971) 301-316.

[10] A. Prékopa, "Stochastic programming models for inventory control and water storage", Colloquia Mathematica Societatis J. Bolyai (North Holland) 7 (1973) 229-245.

[11] A. Prékopa, "Programming under probabilistic constraint with a random technology matrix", Mathematische Operationsforschung und Statistik 5 (1974) 109-116.

[12] A. Prékopa, T. Rapcsák and I. Zsuffa, "Serially linked reservoir system design using stochastic programming", Water Resources Research, to appear.

[13] A. Prékopa, "Logarithmic concave measures and related topics", Proceedings of the International Conference on Stochastic Programming (Oxford, 1974), (Academic Press), to appear.

[14] H. Raiffa and R. Schlaifer, Applied statistical decision theory (M.I.T. Cambridge, Mass., 1961).

[15] J. S. Rustagi, Optimizing methods in statistics (Academic Press, New York, 1971).

[16] J. Todhunter, A history of mathematical theory of probability from the time of Pascal to that of Laplace (Chelsea, New York, 1949).

AP/scr

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER MRC-TSR-1834	2. GOVT ACCESSION NO. Technical	3. RECIPIENT'S CATALOG NUMBER 14 MRC-TSR-1834
4. TITLE (and Subtitle) 6 THE USE OF STOCHASTIC PROGRAMMING FOR THE SOLUTION OF SOME PROBLEMS IN STATISTICS AND PROBABILITY,	9	5. TYPE OF REPORT & PERIOD COVERED Summary Report, ~~no specific reporting period~~ 3
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 András Prékopa		8. CONTRACT OR GRANT NUMBER(s) 15 DAAG29-75-C-0024
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit No. 5 - Mathematical Programming and Operations Research
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709		12. REPORT DATE 11 March 1978 12 22p.
		13. NUMBER OF PAGES 19
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Chance constrained programming
Statistical decision
Testing hypotheses
Dam problem

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The applicability of known stochastic programming models and methods for the solution of problems in classical statistics and probability is shown by a number of examples. These concern testing of hypotheses, constructing of tolerance regions, planning of optimal sampling and the Moran model for the dam.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 221 200 8

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)